SUSCEPTIBILITY LOCUS FOR OSTEOARTHRITIS

Technical Field

This invention relates to the identification of chromosomal regions linked to susceptibility to osteoarthritis using linkage and association analysis.

Background of Invention

Osteoarthritis (OA) is a debilitating disease involving degeneration of the articular cartilage of synovial joints[5,6] Although OA has long been considered an inevitable consequence of ageing, it has become increasingly apparent that OA does have a genetic component. Early-onset forms of the disease are associated with several osteochondrodysplasias, rare diseases involving abnormal bone and cartilage development that are transmitted as Mendelian traits[7]. However, the OA in these conditions is secondary to the main dysplastic phenotype. The common late-onset form of the disease (idiopathic OA) often has no obvious environmental (i.e. injury) or characteristic physical cause and does not demonstrate a clear mode of inheritance.

Over the last 10 years, many genes for single gene or monogenic diseases, which are relatively rare in the population, have been positioned by linkage analysis in families, and localised to a small enough region to allow identification of the gene. The latter sublocalisation and fine mapping can be carried out in single gene rare diseases because recombinations within families define the boundaries of the minimal interval beyond any doubt. In contrast, in common diseases such as osteoarthritis, diabetes or asthma the presence of the disease mutation does not always coincide with the development of the disease: disease susceptibility mutations in common disorders provide risk of developing of the disease, and this risk is usually much less than 100%. Hence, susceptibility genes in common diseases cannot be localised using recombination events within families, unless

tens of thousands of families are available to fine map the locus. Because collections of this size are impractical, investigators are contemplating the use of association mapping, which relies on historical recombination events during the history of the population from which the families came from.

Association mapping has been used in over a dozen examples of rare single gene traits, and particularly in genetically isolated populations such as Finland to fine map disease mutations. Nevertheless, association mapping is fundamentally different from straightforward linkage mapping because even though the degree of association between two markers or a marker and a disease mutation is proportional to the physical distance along the chromosome this relationship can be unpredictable because it is dependent on the allele frequencies of the markers, the history of the population and the age and number of mutations at the disease locus. For rare, highly penetrant single gene diseases there is usually one major founder chromosome in the population under study, making it relatively feasible to locate an interval that is smaller than one that can be defined by standard recombination events within living families. The resolution of this method in monogenic diseases in which there is one main founder chromosome is certainly less than 2cM, and in certain examples the resolution is down to 100 kb of DNA (Hastbacka *et al.* (1994) Cell 78,1-20).

In common diseases like OA, which are caused by a number of genes or polygenes acting together in concert the population frequency of the disease allele may be very high, perhaps exceeding 50%, and there are likely to be several founder chromosomes, all of which impart risk, and not a 100% certainty of disease development. Because association mapping is dependent on unpredictable parameters, and because founder chromosomes will be several and common in frequency in the general population, the task of fine mapping polygenes

is currently one of some controversy, and many doubt the feasibility at all of a systematic genetic approach using a combination of linkage and association mapping. Recently, Risch and Marakandis have provided some mathematical background to the feasibility of association mapping in complex diseases (*Science* 273 1516-1517, 1996) but they did not take into account the effect of multiple founder chromosomes.

It has often been noted in epidemiological studies that there is a female preponderance for OA[5,6]. This may be accounted for by differential effects on the two sexes of environmental factors. However, a Finnish twin study has suggested that genetic susceptibility may be greater in women than men[9]. This result has recently been supported by a segregation analysis[3]. Not only have differences in risk between females and males been reported but it has also been suggested that there are differences in heritability between joint groups[2,10,11]. These differences could be the result of genetic locus heterogeneity between the different joints.

Summary of the Invention

The present inventors have identified a region on chromosome 11q that may harbour a susceptibility locus for OA. This region was identified following a two-stage non-parametric linkage analysis. The first stage involved a random screen of the genome using 272 microsatellite markers in 297 OA families. The second stage was more selective and involved genotyping an additional 184 families for those markers that demonstrated moderate evidence of linkage in stage 1. This revealed three micro satellites on chromosome 11 for whom the evidence for linkage increased as the number of families studied increased. Finer mapping in and around these microsatellites was performed which provided enhanced evidence for linkage and enabled linked regions to be defined. Stratification analysis suggested that the chromosome 11 locus may have differential penetrance between

males and females and between the two different joint groups examined (hips and knees).

The chromosome 11 linkage appears to be accounted for principally by females: the highest single point LOD score obtained was 3.32 (p=0.00005) for marker D11S901 in affected female-only pairs. In male-only pairs the LOD score for D11S901 was 0.04 (p=0.33). The maximum multipoint LOD (MMLS) score in our un-stratified data was 3.15 in-between markers D11S1358 and D11S35 whilst the highest MMLS in our stratified data was in our female/hip pairs with a MMLS of 3.03 in-between markers D11S1314 and D11S916. Linkage disequilibrium analysis identified excess transmission of the 258bp allele (allele 11 as listed in the GDB) of marker D11S937 (5cM proximal to D11S901) in our un-stratified data ($p<0.001$), in our female-only pairs ($0.01>p>0.001$), in our hip-only pairs ($p<0.001$) and in our female/hip pairs ($p<0.001$). This data indicates that the OA susceptibility locus on 11q is located close to D11S937 and that the 258bp allele of D11S937 may be present on a predisposing ancestral chromosome.

Some evidence has also been obtained for linkage with the polymorphic chromosome 6 marker D6S273 and the polymorphic X chromosome marker DXS1068. This may indicate the presence of loci on these chromosomes that also have a role in OA susceptibility. These loci may interact with the locus identified on chromosome 11 and play a role in the differential penetration observed between males and females.

Open reading frames (ORFs) are stretches of genetic sequence which are candidates for being expressed genes. They can be identified by the presence of sequence elements which are characteristic of coding sequence, such as sequence elements from exon-intron boundaries, transcriptional initiation and termination motifs and start and stop codons. Because large amounts of sequence can be screened rapidly for these elements, the identification of ORFs is commonly an initial

step in the discovery of novel genes.

Microsatellite marker loci are designated in this application according to the nomenclature conventional in the field of Human Genetics. This provides a unique designation which specifically and unambiguously identifies each marker locus. Mapping data for any particular marker locus is readily available from conventional sources, such as the Gopher server at the Human Genome Mapping Project Resource Centre (Host = gopher.hgmp.mrc.ac.uk; Port = 70 or URL: gopher://gopher.hgmp.mrc.ac.us:70/ or by anonymous ftp from ftp.hgmp.mrc.uk:/Oxford_Primers).

The present invention relates to chromosomal regions linked to genetic sequences which affect susceptibility to osteoarthritis.

A first aspect of the present invention is a method for identifying individuals susceptible to osteoarthritis comprising obtaining a sample of genomic DNA and detecting the presence or absence of the 258bp allele of D11S937 from chromosome 11.

Another aspect of the present invention is a method for identifying female individuals susceptible to osteoarthritis comprising obtaining a sample of genomic DNA and detecting the presence or absence of the 258bp allele of D11S937 from chromosome 11.

Another aspect of the present invention is a method for identifying individuals susceptible to osteoarthritis of the hip comprising obtaining a sample of genomic DNA and detecting the presence or absence of the 258bp allele of D11S937 from chromosome 11.

A further aspect of the present invention is a method for identifying females susceptible to osteoarthritis of the hip

comprising obtaining a sample of genomic DNA and detecting the presence or absence of the 258bp allele of D11S937 from chromosome 11.

Another aspect of the present invention is a method for isolating genetic loci associated with susceptibility to OA comprising screening a genomic library from an individual who is homozygote for the 258 bp allele of D11S937 and identifying open reading frames in regions adjacent to said allele.

Another aspect of the present invention is a method for isolating genetic loci associated with susceptibility to OA comprising identifying open reading frames in regions adjacent to D11S937 and comparing said open reading frames in individuals carrying a 258 bp allele of D11S937 with said open reading frames in individuals with a non-258 bp allele of D11S937.

Another aspect of the present invention is a method for isolating genetic loci associated with susceptibility to OA comprising screening genomic libraries with the 258 bp allele of D11S937 and identifying open reading frames located within 500 Kb of D11S937, or more preferably within 100 Kb of D11S937, or even more preferably within 50 Kb of D11S937 or most preferably within 10 Kb of D11S937.

Another aspect of the present invention is the use of the 258bp allele of D11S937 as a marker for the identification of loci influencing susceptibility to OA.

Another aspect of the present invention is the use of the 258 bp allele of D11S937 as a marker for the ancestral DNA variant conferring enhanced OA susceptibility .

Another aspect of the present invention is the use of the 258 bp allele of D11S937 as a marker for mapping associated loci

to identify genes associated with OA susceptibility.

Another aspect of the present invention is the use of the 258 bp allele of D11S937 in investigating loci capable of influencing susceptibility to OA.

Another aspect of the present invention is a method for mapping loci which affect susceptibility to OA by comparing genomic regions containing the 258 bp allele of D11S937 with genomic regions containing other alleles of D11S937.

Another aspect of the present invention is a method for determining individual susceptibility to osteoarthritis comprising obtaining sample genomic DNA from siblings, at least two of which have clinical symptoms of osteoarthritis. analysing a region of their genomic DNA comprising the polymorphic marker DS11901, identifying allele sharing between the siblings as defined by a maximum log of the odds (LOD) score of greater than 1 and a p-value of less then 0.25, and determining individual susceptibility to osteoarthritis by reference to the allele sharing.

Another aspect of the present invention is a method for determining individual susceptibility to osteoarthritis comprising obtaining sample genomic DNA from siblings, at least two of which have clinical symptoms of osteoarthritis, analysing a region of their genomic DNA comprising the polymorphic marker DS11901, and additionally analysing one or more of the following; a genomic region comprising the polymorphic marker D6S273 and a genomic region comprising the polymorphic marker DXS1068, identifying allele sharing between the siblings as defined by a maximum log of the odds (LOD) score of greater than 1 and a p-value of less then 0.25, and determining individual susceptibility to osteoarthritis by reference to the allele sharing.

Another aspect of the present invention is a method for

determining individual susceptibility to osteoarthritis comprising obtaining sample genomic DNA from siblings, at least two of which have clinical symptoms of osteoarthritis. analysing a region of their genomic DNA comprising a polymorphic marker and located within 20cM of DS11901, or preferably within 15 cM of DS11901, or more preferably within 10cM of DS11901, even more preferably within 5cM of DS11901 and most preferably within 2cM of DS11901, identifying allele sharing between the siblings as defined by a maximum log of the odds (LOD) score of greater than 1 and a p-value of less then 0.25, and determining individual susceptibility to osteoarthritis by reference to the allele sharing.

Another aspect of the present invention is a method for determining individual susceptibility to osteoarthritis comprising obtaining sample genomic DNA from siblings, at least two of which have clinical symptoms of osteoarthritis. analysing a region of their genomic DNA comprising a polymorphic marker and located within 20cM of DS11901, or preferably within 15 cM of DS11901, or more preferably within 10cM of DS11901, even more preferably within 5cM of DS11901 and most preferably within 2cM of DS11901, and additionally analysing one or more of the following; a genomic region comprising the polymorphic marker D6S273 and a genomic region comprising the polymorphic marker DXS1068, identifying allele sharing between the siblings as defined by a maximum log of the odds (LOD) score of greater than 1 and a p-value of less then 0.25, and determining individual susceptibility to osteoarthritis by reference to the allele sharing

Another aspect of the present invention is a method for determining individual susceptibility to osteoarthritis comprising obtaining sample genomic DNA from siblings, at least two of which have clinical symptoms of osteoarthritis. analysing a region of their genomic DNA comprising one or more of the polymorphic markers D11S901, D11S903, D11S907, identifying allele sharing between the siblings as defined by

a maximum log of the odds (LOD) score of greater than 1 and a p-value of less then 0.25, and determining individual susceptibility to osteoarthritis by reference to the allele sharing.

Another aspect of the present invention is a method for determining individual susceptibility to osteoarthritis comprising obtaining sample genomic DNA from siblings, at least two of which have clinical symptoms of osteoarthritis. analysing a region of their genomic DNA comprising one or more of the polymorphic markers D11S901, D11S903, D11S907, and additionally analysing one or more of the following; a genomic region comprising the polymorphic marker D6S273 and a genomic region comprising the polymorphic marker DXS1068, identifying allele sharing between the siblings as defined by a maximum log of the odds (LOD) score of greater than 1 and a p-value of less then 0.25, and determining individual susceptibility to osteoarthritis by reference to the allele sharing.

Another aspect of the present invention is a method for determining individual susceptibility to osteoarthritis comprising obtaining sample genomic DNA from siblings, at least two of which have clinical symptoms of osteoarthritis. analysing a region of their genomic DNA comprising any one of the polymorphic markers; D2S202, D3S1266, D4S231, D4S415, D6S260, D6S273, D6S286, D6S281, D7S669, D7S530, D11S907, D11S903, D11S901, D17S807, D17S789, DXS1068, identifying allele sharing between the siblings as defined by a maximum log of the odds (LOD) score of greater than 1 and a p-value of less then 0.25, and determining individual susceptibility to osteoarthritis by reference to the allele sharing.
Another aspect of the present invention is a method for determining individual susceptibility to osteoarthritis comprising obtaining sample genomic DNA from siblings, at least two of which have clinical symptoms of osteoarthritis. analysing a region of their genomic DNA comprising any one of

the polymorphic markers D6S273, DXS1068, identifying allele sharing between the siblings as defined by a maximum log of the odds (LOD) score of greater than 1 and a p-value of less then 0.25, and determining individual susceptibility to osteoarthritis by reference to the allele sharing.

Another aspect of the present invention is a method for identifying loci conferring susceptibility to osteoarthritis comprising screening a genomic library with genetic sequence derived from one or more of the following polymorphic markers; D2S202, D3S1266, D4S231, D4S415, D6S260, D6S273, D6S286, D6S281, D7S669, D7S530, D11S907, D11S903, D11S901, D17S807, D17S789, DXS1068 and identifying those isolated clones which additionally contain open reading frames

Another aspect of the present invention is a method for identifying loci conferring susceptibility to osteoarthritis comprising screening a genomic library with genetic sequence derived from one or more of the following polymorphic markers; D2S202, D3S1266, D4S231, D4S415, D6S260, D6S273, D6S286, D6S281, D7S669, D7S530, D11S907, D11S903, D11S901, D17S807, D17S789, DXS1068 and identifying open reading frames located within 500 Kb, more preferably within 100 Kb, even more preferably within 50 Kb or most preferably within 10 Kb of any of these polymorphic markers.

Another aspect of the present invention is a method for isolating genetic loci associated with susceptibility to OA comprising screening a genomic library with sequence derived from the region between polymorphic markers D11S1358 and D11S35 and identifying those isolated clones which additionally contain open reading frames.

Another aspect of the present invention is a method for identifying loci conferring susceptibility to osteoarthritis comprising screening a genomic library with genetic sequence derived from one or more of the polymorphic markers D6S273,

DXS1068 and identifying open reading frames located within 500 Kb, more preferably within 100 Kb, even more preferably within 50 Kb or most preferably within 10 Kb of either of these polymorphic markers.

Description of Drawings

Figure 1 shows a multipoint log of the odds (LOD) analysis on the unstratified data.

Figure 2 shows a multipoint log of the odds (LOD) analysis on stratified data from female sibling pairs only.

Figure 3 shows a multipoint log of the odds (LOD) analysis on stratified data from hip pairs only.

Figure 4 shows a multipoint log of the odds (LOD) analysis on stratified data from female/hip pairs only.

Detailed Description of Invention

Linkage Analysis

The present inventors collected 481 families in which at least two siblings have undergone joint replacement surgery of the hip or knee for severe, end-stage idiopathic OA (Table 1). Due to the late onset of the disease, none of these families contained parents who could participate in the study. In stage 1 272 microsatellite markers were genotyped in 297 of the 481 families. The microsatellites were essentially those used by Reed *et al*[8] with the replacement of certain markers that amplified unreliably. Sixteen markers from stage 1 had evidence of linkage at $p \leq 0.05$ (Table 2). These markers were then genotyped in the remaining 184 families. None of the 16 markers had a $p \leq 0.05$ in this second stage although three had a $p \leq 0.10$: D2S202 ($p=0.07$), D11S903 ($p=0.07$) and D11S901 ($p=0.10$) (Table 2). When the data for stages 1 and 2 were combined and compared to stage 1 only, the combined p-value decreased for 4 of the 16 markers: D2S202 ($p=0.009$ for combined vs. 0.036 for stage 1), D11S907 ($p=0.019$ for combined vs. 0.05 for stage 1), D11S903 ($p=0.004$ for combined vs. 0.017 for stage 1) and D11S901 ($p=0.0003$ for combined vs. 0.0004 for stage 1) increasing the number of

families had therefore increased the evidence for linkage at these 4 markers. The present invention is related to the chromosome 11 markers (D11S901, D11S907 and D11S903) thus identified.

The three chromosome 11 markers that had $p \leq 0.05$ in the combined data set encompass approximately 55cM of the chromosome from 11p13-11q13.2. Most of the chromosome 11 markers from stage 1 were not taken through to stage 2 as they did not reach our linkage criteria of $p \leq 0.05$. Two of these markers immediately flanked the 55cM region whilst one was located within it. These three markers were genotyped in the 184 families of stage 2. In addition, we genotyped 11 new markers located within and around this 55cM region in all 481 families to give us denser coverage of this area of the chromosome. Table 3 lists the results for all 17 chromosome 11 markers that were genotyped in all 481 families. The lowest p-value was still for marker D11S901 (0.0003). A multipoint analysis gave a maximum LOD score (MMLS) of 3.15 distal to D11S901 and in-between markers D11S1358 and D11S35 (Figure 1).

The results were stratified into six categories: those families that were affected female-only pairs (196 families), affected male-only pairs (male and/or female) (54 families), affected female-only pairs who had undergone hip replacement but not knee replacement (female/hip pairs) (132 families) and affected male-only pairs who had undergone hip replacement but not knee replacement (male-hip pairs) (72 families) (Table 1). We did not stratify for female/knee pairs or male/knee pairs as the number of families were too low (21 and 8 respectively) and thus any significant results may simply be the result of stochastic variation.

Stratification revealed that the linkage to chromosome 11 was predominantly accounted for by affected female-only pairs with a single point LOD score of 3.32 (p=0.00005) for marker

D11S901 (Table 4). Five markers immediately flanking D11S901 also supported linkage in female-only pairs at $p \leq 0.05$ with linkage extending from D11S1314 (p=0.006) to D11S1358 (p=0.015), a distance of approximately 20cM. The MMLS for female-only pairs was 2.81 in-between markers D11S901 and D11S1342 (Figure 2). For affected male-only pairs there was no evidence of linkage at D11S901 (p=0.33) or with any of the markers distal or proximal to D11S901 except for marker D11S4089 (p=0.043, 40cM distal to D11S901). There was much greater evidence for linkage in hip-only pairs at D11S901 (p=0.0004) than in knee-only pairs (p=0.093). However, this difference may be due to stochastic variation as there were only 54 knee-only pairs compared to 311 hip-only pairs. The MMLS in hip-only pairs was 2.58, again in-between markers D11S901 and D11S1342 (Figure 3). In female/hips the p-value for D11S901 was 0.001 compared to 0.12 in male/hip pairs. The MMLS in female/hips was 3.03 in-between markers D11S1314 and D11S916 (Figure 4). Overall these results suggest that the linkage to chromosome 11 in our families is restricted to females with OA.

Markers from other chromosomes

Since stratification analysis of chromosome 11 revealed apparent significant differences between the different categories examined we also stratified the 12 other markers that had a $p \leq 0.05$ in stage 1 and which were not on chromosome 11. The majority of these 12 markers will of course represent false positives from our stage 1 analysis: the p-value for each increased when the combined analysis was compared to stage 1 (Table 2). Nevertheless, they cannot all be discounted. Table 5 lists the stratification results for these 12 markers.

Two of the 12 markers are of interest when one considers the results of previous studies: D6S273 and DXS1068.

D6S273 maps close to the HLA complex on chromosome 6p and

Pattrick *et al*[14] have reported association of nodal OA with a specific allele of the HLA-A gene. Furthermore, the COL11A2 gene, which encodes the α2 chain of the cartilage collagen type XI, is tightly linked to the HLA complex. This gene has been linked to and found to harbour causal mutations for the osteochondrodysplasia Stickler syndrome[15]. This syndrome has severe, precocious OA as one of its many phenotypic components.

Leppavuori *et al*[13] have reported linkage of DIP OA to cytogenetic band Xp11.3. DXS1068 also maps to Xp11.3.

Association Analysis

Having established convincing evidence for linkage to chromosome 11, the chromosome 11 markers were tested for association. We used the Transmit software program[16] to test for association with linkage disequilibrium by the transmission disequilibrium test.

For chromosome 11 we analysed the genotyping data without stratification and with the following stratification criteria: female-only pairs, hip-only pairs and female/hip pairs. We tested eight markers: D11S901 plus the seven markers which immediately flanked this marker: D11S1314, D11S916, D11S937, D11S1342, D11S1358, D11S35 and D11S898. These eight markers cover the linkage detected in the un-stratified and the stratified data sets.

In the un-stratified data disequilibrium was positive only for marker D11S937 with a global $\chi^2_{4df}=12.35$ ($0.02>p>0.01$) (Table 6). Disequilibrium was also positive for this marker in affected female-only pairs with global $\chi^2_{4df}=15.19$ ($0.01>p>0.001$), in hip-only pairs with global $\chi^2_{4df}=18.01$ ($0.01.>p>0.001$) and in female/hip pairs with global $\chi^2_{3df}=$ 12.73 ($0.01>p>0.001$). In each case it was the same allele of D11S937 that demonstrated significant distortion in transmission (258bp allele[allele 11] as listed in the GDB):

observed transmission in all families of 254 compared to expected transmission of 233.88 (χ^2_{1df}=11.38, p<0.001), observed transmission in affected female-only pairs of 102 compared to expected transmission of 91.8 (χ^2_{1df}=10.01, 0.01>p>0.001), observed transmission in hip-only pairs of 159 compared to expected transmission of 139.69 (χ^2_{1df}= 17.74,, p<0.001) and observed transmission in female/hip pairs of 65 compared to expected transmission of 55.8 (χ_{1df}= 12.01, p<0.001). Marker D11S937 is only 5cM proximal to D11S901, the marker which has the strongest single-point evidence for linkage in our un-stratified and stratified analysis (Table 4). Our linkage and association results therefore indicate that the OA susceptibility locus on 11q is located close to D11S937 and that the 258bp allele of D11S937 may be present on an ancestral chromosome that harbours a casual DNA variant.

Of the seven other chromosome 11 markers tested, disequilibrium was positive in the stratified data for marker D11S901, but only in the hip-only pairs with a χ^2_{6df} 15.44 (0.02>p>0.01). This disequilibrium was accounted for by the significant ($p \leq 0.05$) distorted transmission of two alleles of D11S901: the 168bp allele (allele 3) and the 170 bp allele (allele 5) as listed in the GDB. The observed transmission in hip-only pairs of the 168bp allele was 384 compared to expected transmission of 371.6 (χ^2_{1df}=4.05, 0.05>p>0.02), whilst the observed transmission of the 170 bp allele was 190 compared to the expected transmission of 180.28 (χ^2_{1df}=4.10, 0.05>p>0.02). Since two alleles are accounting for the positive disequilibrium, it is likely that this disequilibrium represents a false positive.

Methods

Affected sibling-pairs

Families were recruited which contained at least two siblings two had undergone one or more total hip (THR) and/or total knee replacements (TKR) for idiopathic OA. Clinically these

patients are at the severe end of the OA spectrum with advanced radiological changes. The idiopathic diagnosis was supported by clinical, radiological, operative and histological findings: patients who had undergone joint replacement surgery secondary to other factors, such as intra-articular fracture or rheumatoid arthritis, were excluded. Families were ascertained at three centres within the United Kingdom: The Nuffield Orthopaedic Centre in Oxford, the Royal Orthopaedic Hospital in Birmingham and Musgrave Park Hospital in Belfast. Idiopathic OA is typically a late-onset disease and parents of affected siblings are rarely available. Of the 481 families used in the study none contained a parent who was able to participate. We therefore collected siblings who had not undergone joint replacement to assist in the determination of identity-by-descent (IBD) allele transmittance. The 481 families were comprised of 1052 affected individuals plus an additional 302 unaffected siblings (Table 1). 59.3% of the affected individuals were female, 40.7% were male. For each individual, 25ml of venous blood was collected into EDTA tubes and DNA was extracted by conventional techniques.

Markers and Genotyping

Our initial screening panel of 272 microsatellite markers was essentially the panel used by Reed *et al*[8]. The additional microsatellite markers used to provide a denser coverage of chromosome 11 were obtained from the GDB or from the ABI Prism Linkage Mapping Set (Version 2). The markers were amplified using conventional conditions with either the forward or the reverse primer in a PCR pair fluorescently labelled. The amplification products were electrophoresed through 6% acrylamide using an Applied Biosystems 377 Automated DNA Sequencer[B]. Alleles were sized using Applied Biosystems Genescan version 2.0.2. and Genotyper version 1.1 software.

Linkage and Linkage disequilibrium analysis

Initially error checking procedures were employed for all families for each marker. After identification of straightforward mis-inheritances, more subtle transmission errors were detected using the PedCheck program[17]. The entire family data set was tested with Relative which is able to produce a probability calculation (based on 50 or more unlinked markers) that full sibships are in fact half sibs or even unrelated (due to unknown adoption of laboratory error). All 481 families used in the study successfully progressed through these checks. In addition markers were checked for having excess homozygotes, based on their allele frequencies and heterozygosities. Markers shown to be unreliable were eliminated from the study. Finally the marker data were haplotyped for each chromosome using Simwalk2. This checks for areas on the chromosome where excessive recombination events may alert us to genotyping errors or mis-assignment of a marker position.

Non parametric linkage analysis was performed using the SIBPAIR module of the ANALYZE package[18]. This module is able to use data from siblings to determine identity-by-descent (IBD) allele transmittance. In this way it is appropriate to our study since we were unable to collect parents of our affected siblings. In the linkage analysis, siblings who had not undergone joint replacement were given a clinical status of unknown. The SIBPAIR module produces a singlepoint LOD score and its asymptotic P-value. Allele frequencies were calculated from the input data using either GAS or Downfreq (part of the ANALYZE package). Subsequent multipoint analyses was performed using ASPEX which calculates its own allele frequencies from the data set, using a maximum likelihood method, and employs marker information across the chromosome simultanenously[19]. ASPEX produces maximum LOD score (MLS) under an additive model. In addition it produces an exclusion map along the entire chromosome based on a fixed value for λs.

We tested for linkage disequilibrium by the transmission disequilibrium test (TDT) using the TRANSMIT software program[16]. Alleles with a frequency <0.1 were pooled. A global χ^2 statistic was computed for each microsatellite. If a microsatellite was significant a χ^2 statistic was then computed for its individual alleles.

Stratification

We stratified for sex, for joint replaced (hip or knee) and for sex combined with joint replaced.

For those families with more than two affected siblings and in which the siblings were not of the same sex, the affected sibling of opposite sex to a same pair was given an affected status of unknown in the linkage analysis. In this way we were stratifying for sex whilst not excluding siblings who could be used to assist in the determination of identity-by-descent (IBD) allele transmittance.

A hip-only pair were siblings who had each undergone joint replacement of the hip only (mono or bi-lateral) whilst a knee-only pair had undergone joint replacement of the knee only (mono or bi-lateral). If an affected pair was composed of one sibling who had undergone joint replacement of one joint type only (hip or knee) whilst their affected sibling had undergone joint replacement of the hip and knee then that pair were excluded. For an affected trio, if a pair of the siblings had both undergone joint replacement of the same joint type only (hip or knee) whilst the third sibling had undergone both hip and knee replacement, then the concordant pair were used in the stratification study whilst the third sibling was given an unaffected status in the linkage analysis. Again, we were attempting to maximise our determination of IBD allele transmittance.

Table 1(a)

Table 1 a) The families used in stages 1 and 2 together with the combined total figure. Also listed are the pairs concordant for different stratification criteria. b) A list of the individuals in the study.

a)

Families

	Stage 1	Stage 2	Total
Families	297	184	481
sibling pair	265	150	415
sibling trio	23	27	50
sibling quattro	7	5	12
other[a]	2	2	4
Female only	132	64	196
pair	121	59	180
trio	11	5	16
Male only	61	41	102
pair	57	36	93
trio	3	5	8
other[a]	1	0	1
Hip only	194	117	311
pair	170	96	266
trio	18	17	35
quattro	4	2	6
other[a]	2	2	4
Knee only	34	20	54
pair	33	19	52
trio	1	1	2
Female/hip only	85	47	132
pair	77	46	123
trio	8	1	9
Female/knee only	16	5	21
pair	16	5	21
Male/hip only	45	27	72
pair	41	25	66
trio	3	2	5
other[a]	1	0	1
Male/knee only	4	4	8
pair	4	4	8

Table 1(b)

b)

Individuals

	Stage 1	Stage 2	Total
Affected individuals	641	411	1052
Female	394	230	624
Male	247	181	428
Hip only	479	309	788
Knee only	121	77	198
Hip & Knee	41	25	66
Female/hip only	287	172	459
Female/knee only	77	42	119
Female/hip & knee	30	16	46
Males/hip only	192	137	329
Males/knee only	44	35	79
Males/hip & knee	11	9	20
Additional siblings[b]	211	91	302
Female	107	49	156
Male	104	42	146

[a]Other relative pairs such as cousins, uncles, aunts.
[b]Since our families lack parents, siblings who had not undergone joint replacement surgery were collected to assist in the determination of parental genotypes. These siblings were given an "unknown" clinical status in the linkage analysis.

Table 2 LOD scores and p-values for all markers that had a p≤0.05 in screen 1, for these markers in screen 2 and for screens 1 and 2 combined (* = p≤0.05).

	STAGE 1		STAGE 2		COMBINED	
Marker	p-value	LOD	p-value	LOD	p-value	LOD
D2S202	0.036*	0.70	0.07	0.49	0.009*	1.21
D3S1266	0.017*	0.96	0.5	0.00	0.082	0.42
D4S231	0.040*	0.67	0.5	0.00	0.33	0.04
D4S415	0.018*	0.95	0.33	0.04	0.025*	0.83
D6S260	0.050*	0.58	0.5	0.00	0.13	0.29
D6S273	0.016*	0.98	0.5	0.00	0.077	0.44
D6S286	0.030*	0.77	0.5	0.00	0.081	0.42
D6S281	0.046*	0.61	0.45	0.00	0.062	0.52
D7S669	0.018*	0.94	0.25	0.10	0.021*	0.90
D7S530	0.006*	1.33	0.41	0.01	0.013*	1.09
D11S907	0.050*	0.58	0.12	0.31	0.019*	0.92
D11S903	0.017*	0.97	0.07	0.49	0.004*	1.45
D11S901	0.0004*	2.45	0.10	0.37	0.0003*	2.51
D17S807	0.014*	1.03	0.5	0.00	0.15	0.24
D17S789	0.010*	1.16	0.5	0.00	0.071	0.47
DXS1068	0.020*	0.84	0.5	0.00	0.10	0.35

Table 3 LOD scores and p-value for stages 1 and 2 combined for the chromosome 11 markers.

Marker	cM from 11p telomere	p-value	LOD
D11S4046	4	0.36	0.03
D11S907	51	0.019*	0.92
D11S903	64	0.004*	1.45
D11S4191	69	0.080	0.43
D11S1883	73	0.50	0.00
D11S1314	82	0.004*	1.50
D11S916	85	0.008*	1.22
D11S937	89	0.060	0.55
D11S901	94	0.0003*	2.51
D11S1342	101	0.026*	0.81
D11S1358	102	0.050*	0.56
D11S35	110	0.090	0.37
D11S898	110	0.050*	0.56
D11S908	121	0.50	0.00
D11S925	133	0.08	0.41
D11S4089	134	0.45	0.00
D11S1320	157	0.50	0.00

Table 4 Stratification of stages 1 and 2 combined of the chromosome 11 markers for affected female-only pairs, affected male-only pairs, hip-only pairs, knee-only pairs, affected females with hip-only disease and affected males with hip-only disease.

		FEMALES		MALES		HIPS		KNEES		FEMALE/HIPS		MALE/HIPS	
Marker	cM	p-value	LOD	p-value	LOD	p-value	LOD	p-value	LOD	p-value	LOD	p-value	LOD
D11S4046	4	0.50	0.00	0.025*	0.83	0.28	0.07	0.29	0.07	0.50	0.00	0.0048*	1.45
D11S907	51	0.11	0.32	0.50	0.00	0.050*	0.59	0.098	0.36	0.15	0.24	0.36	0.03
D11S903	64	0.50	0.00	0.34	0.04	0.086	0.40	0.30	0.05	0.50	0.00	0.50	0.00
D11S4191	69	0.068	0.48	0.23	0.12	0.49	0.00	0.32	0.05	0.18	0.18	0.50	0.00
D11S1883	73	0.38	0.02	0.50	0.00	0.50	0.00	0.50	0.00	0.44	0.01	0.50	0.00
D11S1314	82	0.006*	1.37	0.50	0.00	0.002*	1.82	0.087	0.40	0.002*	1.73	0.50	0.00
D11S916	85	0.015	1.02	0.50	0.00	0.0009*	2.12	0.50	0.00	0.004*	1.55	0.32	0.05
D11S937	89	0.015	1.03	0.50	0.00	0.015*	1.02	0.50	0.00	0.019*	0.94	0.50	0.00
D11S901	94	0.00005*	3.32	0.33	0.04	0.0004*	2.45	0.093	0.38	0.001*	2.04	0.12	0.31
D11S1342	101	0.019*	0.94	0.30	0.06	0.086	0.41	0.044*	0.63	0.16	0.21	0.13	0.28
D11S1358	102	0.015*	1.02	0.32	0.05	0.13	0.26	0.34	0.04	0.021*	0.89	0.32	0.05
D11S35	110	0.083	0.42	0.066	0.49	0.10	0.35	0.50	0.00	0.038*	0.68	0.077	0.44
D11S898	110	0.063	0.51	0.11	0.33	0.037*	0.69	0.46	0.00	0.029*	0.78	0.16	0.22
D11S908	121	0.50	0.00	0.13	0.29	0.50	0.00	0.15	0.23	0.30	0.06	0.26	0.09
D11S925	133	0.083	0.42	0.08	0.44	0.17	0.20	0.50	0.00	0.10	0.37	0.24	0.11
D11S4089	134	0.5	0.00	0.043*	0.64	0.37	0.02	0.50	0.00	0.50	0.00	0.020*	0.91
D11S1320	157	0.5	0.00	0.50	0.00	0.50	0.00	0.49	0.00	0.50	0.00	0.50	0.00

Table 5 Stratification of stages 1 and 2 combined of the 12 markers that had a p≤0.05 for stage 1 (excluding the chromosome 2 and 11 markers) for affected female-only pairs, affected male-only pairs, hip-only pairs, knee-only pairs, affected females with hip-only disease and affected males with hip-only disease.

	All Families		Females		Males		Hips		Knees		Female/Hips		Male/Hips	
Marker	p	LS	p	LS	p	LS	p	LS	p	LS	p	LS	p	LS
D3S1266	0.082	0.42	0.0060*	1.37	0.25	0.09	0.29	0.07	0.0035*	1.58	0.0067*	1.33	0.42	0.01
D4S231	0.33	0.04	0.045*	0.62	0.16	0.21	0.066	0.49	0.10	0.35	0.0052*	1.42	0.26	0.09
D4S415	0.025*	0.83	0.15	0.22	0.11	0.34	0.20	0.16	0.12	0.30	0.50	0.00	0.24	0.11
D6S260	0.13	0.29	0.16	0.22	0.50	0.00	0.19	0.17	0.50	0.00	0.50	0.00	0.33	0.04
D6S273	0.077	0.44	0.0034*	1.59	0.50	0.00	0.18	0.18	0.044*	0.63	0.0053*	1.42	0.5	0.00
D6S286	0.081	0.42	0.13	0.27	0.35	0.03	0.016*	0.99	0.50	0.00	0.058	0.54	0.36	0.03
D6S281	0.062	0.52	0.37	0.03	0.50	0.00	0.11	0.33	0.40	0.01	0.17	0.20	0.50	0.00
D7S669	0.021*	0.90	0.071	0.47	0.081	0.42	0.010*	1.16	0.39	0.02	0.032*	0.74	0.16	0.22
D7S530	0.013*	1.09	0.090	0.39	0.15	0.24	0.034*	0.72	0.17	0.19	0.18	0.18	0.29	0.07
D17S807	0.15	0.24	0.32	0.05	0.5	0.00	0.12	0.30	0.50	0.00	0.18	0.18	0.25	0.10
D17S789	0.071	0.47	0.12	0.29	0.39	0.02	0.25	0.10	0.021*	0.90	0.28	0.07	0.20	0.16
DXS1068	0.10	0.35	0.20	0.16	0.049*	0.60	0.19	0.17	0.26	0.09	0.50	0.00	0.02*	0.90

Table 6
TRANSMIT of Chromosome 11.
$\chi 2$ analysis
(figures in brackets = degrees of freedom)

Locus	All data	Females-only	Hips-only	Female/hips
D11S1314	(4) 3.53	(4) 3.51	(4) 5.8	(4) 7.05
D11S916	(3) 0.67	(3) 1.79	(3) 0.51	(4) 4.81
D11S937	(4) 12.35*	(4) 15.19*	(4) 18.01*	(3) 12.73*
D11S901	(6) 8.21	(6) 7.85	(6) 15.44*	(6) 10.37
D11S1342	(3) 0.55	(3) 4.67	(3) 2.01	(3) 6.91
D11S1358	(4) 5.02	(4) 2.58	(4) 8.21	(4) 2.66
D11S35	(5) 4.39	(5) 1.38	(4) 3.34	(5) 3.18
D11S898	(5) 5.46	(5) 3.27	(4) 3.87	(5) 4.21

References

1. Spector T.D., Cicuttini F., Baker J., Loughlin J. & Hart D. Genetic influences on osteoarthritis in women: a twin study. Br. Med. J. 312, 940-943 (1996).

2. Chitnavis J., Sinsheimer J.S., Clipsham K., Loughlin J., Sykes B., Burge P.D. & Carr A.J. Genetic influences in end-stage osteoarthritis. Sibling risks of hip and knee replacement for idiopathic osteoarthritis, J. Bone. Joint. Surge 79, 660-664 (1997).

3. Felson D.T., Couropmitree N.N., Chaisson C.E., Hannan M.T., Zhang Y., McAlindon T.E., LaValley M., Levy D. & Myers R.H. Evidence for a Mendelian gene in a segregation analysis of generalized radiographic osteoarthritis. Arth. Rheum. 41, 1064-1071 (1998).

4. Hirsch R., Lethbridge-Cejku M., Hanson R., Scott W.W., Reichle R., Plato C.C., Tobin J.D. & Hochberg M.C. Familial aggregation of osteoarthritis. Arth. Rheum. 41, 1227-1232 (1998).

5. Creamer P. & Hochberg M.C. Osteoarthritis. Lancet 350, 503-508 (1997).

6. Jones A. & Doherty M. ABC of Rheumatology. Osteoarthritis. Br. Med. J. 310, 457-460 (1995).

7. Horton W.A. Evolution of the bone dysplasia family. Am. J. Med. Genet. 63, 4-6 (1996).

8. Reed P.W., Davies J.L., Copeman J.B., et al. Chromosome-specific microsatellite sets for fluorescence-based, semi-automated genome mapping. Nature Genet. 7, 390-395 (1994).

9. Kaprio J., Kujala U.M., Peltonen L. & Koskenvuo M. Genetic liability to osteoarthritis may be greater in women

than men. Br. Med. J. 313, 232, (1996).

10. Lindberg H. Prevalence of primary coxarthrosis in siblings of patients with primary coxarthrosis. Clin. Orthop. 203, 273-275 (1986).

11. Cooper C., McAlindon T., Snow S., et al. Mechanical and constitutional risk factors for symptomatic knee osteoarthritis: differences between medial tibio-femoral and patello-femoral disease. J. Rheumatol. 21, 307-313 (1994).

12. Wright G.D., Hughes A.E., Regan M. & Doherty M. Association of two loci on chromosome 2q with nodal osteoarthritis. Ann. Rheum. Dis. 55, 317-319 (1996).

13. Leppavuori J.K., Kujala U.M., Kaprio J., Nissila M., Helliovaara M., Kinnunen J., Koskenvuo M. & Peltonen L. Genome screen for predisposing loci of distal interphalangeal joint osteoarthritis. Am. J. Hum. Genet. Suppl. 63, A1715 (1998).

14. Pattrick M., Manhire A., Ward A.M. & Doherty M. HLA-A, B antigens and αI-antitrypsin phenotypes in nodal generalised osteoarthritis and erosive osteoarthritis. Ann. Rheum. Dis. 48, 470-475 (1989).

15. Mundlos S. & Olsen B.R. Heritable diseases of the skeleton. Part II: Molecular insights into skeletal development-matrix components and their homeostasis. FASEB J. 11, 227-233 (1977).

16. Clayton D. Transmit, http://www.mrc-bsu.cam.ac.uk/pub/methodology/genetics/transmit (1997).

17. O'Connell J.R. & Weeks D.E. PedCheck: a program for identifying marker typing incompatibilities in linkage

analysis. Am. J. Hum. Genet. Suppl. 60, A288 (1997).

18. Terwilliger J. Program SIBPAIR: sibpair analysis on nuclear families.
1996;ftp://linkage.cpmc.columbia.edu.

19. Hinds D. & Risch N. The ASPEX package: affected sibpair mapping.
1996;ftp://lahmed.stanford.edu/pub/aspex.